Pre-training a model on a **large** dataset with **YOLOv8**, and **retraining** this model

on tiny datasets, it effectively detects

specific features previously undetected

A Hierarchical Approach to Pest Detection Using Few Shot Learners and Small Datasets

INTRO

- Advances in machine learning offer new solutions to the issue of pest detection.
- Pests prevalent in South Africa, such as the Sirex and Gonipterus, are underrepresented in existing models and data sets.
- We address this gap by training a model to detect the *Sirex* and *Gonipterus* pests and identify the underlying attributes of insects.

METHODS

• A YOLOv8n model was retrained using the IP102 public dataset containing 19 000 box annotated

RESULTS



More Information Streamlit App:

This is a test deployment of an image detection model created in conjunction with the F Biological Institute (FABI) at the University of Pretoria (UP). The model is an augmentati architecture, trained for the identification of the Sirex and Gonipterus pests often encou South African forestry industry. Please see the project Github for additional information	orestry and on of the YOLOv ntered in the
Show instructions	
Please select a model to use	
YOLO v8 FABI	
Use sample images	
Upload An Image	
Drag and drop file here Limit 200MB per file • PNG, JPEG, JPG	Browse files
Gonipterus_DSC_4204.JPG 6.0MB	×
Select Model Confidence	
40	
At a 40.0% confidence threshold, the model identifies 1 object(s): - gonipterus with 79.13% confidence.	

IP102 Model Performance:



images of insects prevalent in Asia.

- This model was then augmented with data from the Forestry and Biological Institute (FABI), containing 142 Gonipterus and 42 Sirex photos.
- Annotating specific features, 7 models were retrained, focusing on detecting:
 - The Sirex and Gonipterus life stages
 - Any insect
 - Flying insects
 - Larvae
 - Sirex larvae
 - Sirex adults
 - Wings

Marc Gagiano, Stephan Kersop



DISCUSSION

- Training an object detector on tiny datasets are challenging.
- Part of the challenge is solved by pre-training on a larger public datasets
- The solution is further enhanced by annotating specific features on various images and retraining individual models for each of these features
- The features one model misses will be picked up by one of the other models
- Identifying one single feature already reduces time spent on classification tasks



Lifestage Model Performance:



Segmentation Model:







Department of Computer Science

Faculty of Engineering, Built Environment and Information Technology Fakulteit Ingenieurswese, Bou-omgewing en Inligtingtegnologie / Lefapha la Boetšenere, Tikologo ya Kago le Theknolotši ya Tshedimošo

Capstone Project - MIT 808

Course Coordinators: Dr. Vukosi Marivate (vukosi.marivate@cs.up.ac.za) Abiodun Modupe (abiodun.modupe@cs.up.ac.za)